AMENDMENTS TO THE CLAIMS

What is claimed is:

1. (Currently Amended) A method for detecting transmission of potentially unwanted e-mail messages, comprising:

receiving a plurality of e-mail messages;

processing the e-mail messages by removing HTML comments and HTML tags from the e-mail messages;

generating hash values, as generated hash values, based on ~~one or more~~a plurality of portions of each message body of the plurality of e-mail messages that have been processed by removing the HTML comments and the HTML tags, such that each message body of each of the e-mail messages has a corresponding plurality of generated hash values;

~~determining whether~~counting a number of the generated hash values corresponding to the message body associated with one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with prior e-mail messages; and

utilizing a settable score-related threshold, determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message, the determination being based, at least in part, on the number of the generated hash values corresponding to the message body associated with one of the e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages~~when one or more of the generated hash values associated with the one of the plurality of e-mail messages match one or more of the hash values associated with the prior e-mail messages~~.

2. (Original) The method of claim 1, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of variable-sized blocks of a main text of the plurality of e-mail messages.

3. (Original) The method of claim 1, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of fixed-sized blocks of a main text of the plurality of e-mail messages.

4. (Original) The method of claim 1, wherein the generating hash values includes:

performing a plurality of hashes on a main text of the plurality of e-mail messages using a plurality of different hash functions.

5. (Original) The method of claim 1, wherein the generating hash values includes:

performing a plurality of hashes on a main text of the plurality of e-mail messages using a same hash function.

6. (Original) The method of claim 1, wherein the generating hash values includes:

attempting to expand an attachment of the plurality of e-mail messages, and hashing the attachment after attempting to expand the attachment.

7. (Original) The method of claim 1, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of variable-sized blocks of an attachment of the plurality of e-mail messages.

8. (Original) The method of claim 1, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of fixed-sized blocks of an attachment of the plurality of e-mail messages.

9. (Original) The method of claim 1, wherein the generating hash values includes:

performing a plurality of hashes on an attachment of the plurality of e-mail messages using a plurality of different hash functions.

10. (Original) The method of claim 1, wherein the generating hash values includes:

performing a plurality of hashes on an attachment of the plurality of e-mail messages using a same hash function.

11. (Original) The method of claim 1, further comprising:

comparing the generated hash values to hash values corresponding to known unwanted e-mails.

12. (Original) The method of claim 11, wherein the known unwanted e-mails include at least one of e-mails containing a virus, e-mails containing a worm, and unsolicited commercial e-mails.

13. (Original) The method of claim 1, wherein the generating hash values includes:

hashing at least one of a main text and an attachment to generate one or more first hash values, and

hashing a concatenation of first and second header fields to generate a second hash value.

14. (Original) The method of claim 13, wherein the first and second header fields include a From header field and a To header field.

15. (Currently Amended) The method of claim 13, wherein the ~~determining whether~~counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages includes:

determining a first suspicion count based on a number of the hash values associated with the prior e-mail messages that match the one or more first hash values, and

determining a second suspicion count based on a number of the hash values associated with the prior e-mail messages that match the second hash value.

16. (Currently Amended) The method of claim 15, wherein the determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message includes:

determining that the one of the plurality of e-mail messages is a potentially unwanted email message when the first suspicion count is ~~significantly~~ higher than the second suspicion count.

17. (Original) The method of claim 1, further comprising:

taking remedial action when the one of the plurality of e-mail messages is a potentially unwanted e-mail message, the taking remedial action including at least one of:

discarding the one of the plurality of e-mail messages,

bouncing the one of the plurality of e-mail messages,

marking the one of the plurality of e-mail messages with a warning,

subjecting the one of the plurality of e-mail messages to a virus or worm detection process,

creating a notification message, and

generating a suspicion score for the one of the plurality of e-mail messages and using the suspicion score to identify further processing for the one of the plurality of e-mail messages.

18. (Currently Amended) The method of claim 1, further comprising:

generating a suspicion score for the plurality of e-mail messages based on a result of the ~~determination of whether~~counting of the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages; and

taking remedial action when the one of the plurality of e-mail messages is a potentially unwanted e-mail message, the taking remedial action including:

determining whether a newly received e-mail message exceeds a mail quota,

identifying an earlier-received e-mail message with a highest suspicion score,

determining whether the suspicion score of the newly received e-mail message is lower than the suspicion score of the earlier-received e-mail message when the newly received e-mail message exceeds the mail quota,

deleting the earlier-received e-mail message when the suspicion score of the newly received e-mail message is lower than the suspicion score of the earlier-received e-mail message, and

storing the newly received e-mail message.

19. (Currently Amended) The method of claim 1, wherein the generating hash values and the ~~determining whether~~counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages are performed incrementally as the plurality of e-mail messages are being received.

20. (Currently Amended) The method of claim 19, further comprising:

generating a suspicion score for the plurality of e-mail messages based on a result of the ~~determination of whether~~counting of the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages; and

taking remedial action when the suspicion score of an e-mail message of the plurality of e-mail messages is above a threshold, the taking remedial action including rejecting the e-mail message.

21. (Original) The method of claim 20, wherein the rejecting occurs before the e-mail message is completely received.

22. (Original) The method of claim 1, further comprising:

comparing the generated hash values to known legitimate mailing lists; and

passing the plurality of e-mail messages without further examination when the generated hash values match one or more of the known legitimate mailing lists.

23. (Original) The method of claim 22, wherein the comparing the generated hash values includes:

determining whether the plurality of e-mail messages originated from the known legitimate mailing lists.

24. (Original) The method of claim 1, wherein the generating hash values includes:

hashing a main text to generate a first hash value, and

hashing sender-related header fields to generate one or more second hash values.

25. (Original) The method of claim 24, wherein the sender-related header fields include at least one of a From header field, a Sender header field, and a Reply-To header field.

26. (Currently Amended) The method of claim 24, wherein the ~~determining whether~~counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages includes:

determining a first suspicion count based on a number of the hash values associated with the prior e-mail messages that match the first hash value, and

determining one or more second suspicion counts based on a number of the hash values associated with the prior e-mail messages that match the one or more second hash values.

27. (Original) The method of claim 26, wherein the determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message includes:

determining that the one of the plurality of e-mail messages is a potentially unwanted e-mail message when the first suspicion count is higher than the one or more second suspicion counts.

28. (Original) The method of claim 1, wherein the generating hash values includes:

hashing a main text of the plurality of e-mail messages to generate a main text hash, and

hashing at least one header field of the plurality of e-mail messages to generate at least one header hash.

29. (Currently Amended) The method of claim 28, wherein the ~~determining whether~~counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages includes:

determining whether the main text hash matches a ~~substantially~~ higher number of the hash values associated with the prior e-mail messages than the at least one header hash; and

wherein the determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message includes:

determining that the one of the plurality of e-mail messages is a potentially unwanted e-mail message when the main text hash matches a ~~substantially~~ higher number of the hash values associated with the prior e-mail messages than the at least one header hash.

30.-66. (Cancelled)

67. (New) The method of claim 1, wherein the processing the e-mail messages by removing the HTML comments and the HTML tags from the e-mail messages occurs before the generating the hash values.

68. (New) The method of claim 1, wherein the processing the e-mail messages by removing the HTML comments and the HTML tags from the e-mail messages occurs in parallel with the generating the hash values.

69. (New) The method of claim 1, wherein the portions of each message body of the plurality of e-mail messages that have been processed by removing the HTML comments and the HTML tags include blocks.

70. (New) A computer program product embodied on a tangible computer readable medium, comprising:

computer code for receiving a plurality of e-mail messages;

computer code for processing the e-mail messages by removing HTML comments and HTML tags from the e-mail messages;

computer code for generating hash values, as generated hash values, based on a plurality of portions of each message body of the plurality of e-mail messages that have been processed by removing the HTML comments and the HTML tags, such that each message body of each of the e-mail messages has a corresponding plurality of generated hash values;

computer code for counting a number of the generated hash values corresponding to the message body associated with one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with at least one prior e-mail message; and

computer code for, utilizing a settable score-related threshold, determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message, the determination being based, at least in part, on the number of the generated hash values corresponding to the message body associated with one of the e-mail messages that match the hash values corresponding to the message body associated with the at least one prior e-mail message.

71. (New) The computer program product of claim 70, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of variable-sized blocks of a main text of the plurality of e-mail messages.

72. (New) The computer program product of claim 70, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of fixed-sized blocks of a main text of the plurality of e-mail messages.

73. (New) The computer program product of claim 70, wherein the generating hash values includes:

performing a plurality of hashes on a main text of the plurality of e-mail messages using a plurality of different hash functions.

74. (New) The computer program product of claim 70, wherein the generating hash values includes:

performing a plurality of hashes on a main text of the plurality of e-mail messages using a same hash function.

75. (New) The computer program product of claim 70, wherein the generating hash values includes:

attempting to expand an attachment of the plurality of e-mail messages, and

hashing the attachment after attempting to expand the attachment.

76. (New) The computer program product of claim 70, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of variable-sized blocks of an attachment of the plurality of e-mail messages.

77. (New) The computer program product of claim 70, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of fixed-sized blocks of an attachment of the plurality of e-mail messages.

78. (New) The computer program product of claim 70, wherein the generating hash values includes:

performing a plurality of hashes on an attachment of the plurality of e-mail messages using a plurality of different hash functions.

79. (New) The computer program product of claim 70, wherein the generating hash values includes:

performing a plurality of hashes on an attachment of the plurality of e-mail messages using a same hash function.

80. (New) The computer program product of claim 70, further comprising:

comparing the generated hash values to hash values corresponding to known unwanted e-mails.

81. (New) The computer program product of claim 80, wherein the known unwanted e-mails include at least one of e-mails containing a virus, e-mails containing a worm, and unsolicited commercial e-mails.

82. (New) The computer program product of claim 70, wherein the generating hash values includes:

hashing at least one of a main text and an attachment to generate one or more first hash values, and

hashing a concatenation of first and second header fields to generate a second hash value.

83. (New) The computer program product of claim 82, wherein the first and second header fields include a From header field and a To header field.

84. (New) The computer program product of claim 82, wherein the counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the at least one prior e-mail message includes:

determining a first suspicion count based on a number of the hash values associated with the at least one prior e-mail message that match the one or more first hash values, and

determining a second suspicion count based on a number of the hash values associated with the at least one prior e-mail message that match the second hash value.

85. (New) The computer program product of claim 84, wherein the determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message includes:

determining that the one of the plurality of e-mail messages is a potentially unwanted email message when the first suspicion count is higher than the second suspicion count.

86. (New) The computer program product of claim 70, further comprising:

taking remedial action when the one of the plurality of e-mail messages is a potentially unwanted e-mail message, the taking remedial action including at least one of:

discarding the one of the plurality of e-mail messages,

bouncing the one of the plurality of e-mail messages,

marking the one of the plurality of e-mail messages with a warning,

subjecting the one of the plurality of e-mail messages to a virus or worm detection process,

creating a notification message, and

generating a suspicion score for the one of the plurality of e-mail messages and using the suspicion score to identify further processing for the one of the plurality of e-mail messages.

87. (New) The computer program product of claim 70, further comprising:

generating a suspicion score for the plurality of e-mail messages based on a result of the counting of the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the at least one prior e-mail message; and

taking remedial action when the one of the plurality of e-mail messages is a potentially unwanted e-mail message, the taking remedial action including:

determining whether a newly received e-mail message exceeds a mail quota,

identifying an earlier-received e-mail message with a highest suspicion score,

determining whether the suspicion score of the newly received e-mail message is lower than the suspicion score of the earlier-received e-mail message when the newly received e-mail message exceeds the mail quota,

deleting the earlier-received e-mail message when the suspicion score of the newly received e-mail message is lower than the suspicion score of the earlier-received e-mail message, and

storing the newly received e-mail message.

88. (New) The computer program product of claim 70, wherein the generating hash values and the counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the at least one prior e-mail message are performed incrementally as the plurality of e-mail messages are being received.

89. (New) The computer program product of claim 88, further comprising:

generating a suspicion score for the plurality of e-mail messages based on a result of the counting of the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash

values corresponding to the message body associated with the at least one prior e-mail message; and

taking remedial action when the suspicion score of an e-mail message of the plurality of e-mail messages is above a threshold, the taking remedial action including rejecting the e-mail message.

90. (New) The computer program product of claim 89, wherein the rejecting occurs before the e-mail message is completely received.

91. (New) The computer program product of claim 70, further comprising:

comparing the generated hash values to known legitimate mailing lists; and

passing the plurality of e-mail messages without further examination when the generated hash values match one or more of the known legitimate mailing lists.

92. (New) The computer program product of claim 91, wherein the comparing the generated hash values includes:

determining whether the plurality of e-mail messages originated from the known legitimate mailing lists.

93. (New) The computer program product of claim 70, wherein the generating hash values includes:

hashing a main text to generate a first hash value, and

hashing sender-related header fields to generate one or more second hash values.

94. (New) The computer program product of claim 93, wherein the sender-related header fields include at least one of a From header field, a Sender header field, and a Reply-To header field.

95. (New) The computer program product of claim 93, wherein the counting the number of the generated hash values corresponding to the message body associated with

the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the at least one prior e-mail message includes:

determining a first suspicion count based on a number of the hash values associated with the at least one prior e-mail message that match the first hash value, and

determining one or more second suspicion counts based on a number of the hash values associated with the at least one prior e-mail message that match the one or more second hash values.

96. (New) The computer program product of claim 95, wherein the determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message includes:

determining that the one of the plurality of e-mail messages is a potentially unwanted e-mail message when the first suspicion count is higher than the one or more second suspicion counts.

97. (New) The computer program product of claim 70, wherein the generating hash values includes:

hashing a main text of the plurality of e-mail messages to generate a main text hash, and

hashing at least one header field of the plurality of e-mail messages to generate at least one header hash.

98. (New) The computer program product of claim 97, wherein the counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the at least one prior e-mail message includes:

determining whether the main text hash matches a higher number of the hash values associated with the at least one prior e-mail message than the at least one header hash; and

wherein the determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message includes:

determining that the one of the plurality of e-mail messages is a potentially unwanted e-mail message when the main text hash matches a higher number of the hash values associated with the at least one prior e-mail message than the at least one header hash.

99. (New) The computer program product of claim 70, wherein the processing the e-mail messages by removing the HTML comments and the HTML tags from the e-mail messages occurs before the generating the hash values.

100. (New) The computer program product of claim 70, wherein the processing the e-mail messages by removing the HTML comments and the HTML tags from the e-mail messages occurs in parallel with the generating the hash values.

101. (New) The computer program product of claim 70, wherein the portions of each message body of the plurality of e-mail messages that have been processed by removing the HTML comments and the HTML tags include blocks.

102. (New) A system including a tangible computer readable medium, comprising:

logic for receiving a plurality of e-mail messages;

logic for processing the e-mail messages by removing HTML comments and HTML tags from the e-mail messages;

logic for generating hash values, as generated hash values, based on a plurality of portions of each message body of the plurality of e-mail messages that have been processed by removing the HTML comments and the HTML tags, such that each message body of each of the e-mail messages has a corresponding plurality of generated hash values;

logic for counting a number of the generated hash values corresponding to the message body associated with one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with prior e-mail messages; and

logic for, utilizing a settable score-related threshold, determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message, the determination

being based, at least in part, on the number of the generated the hash values corresponding to the message body associated with one of the e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages.

103. (New) The system of claim 102, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of variable-sized blocks of a main text of the plurality of e-mail messages.

104. (New) The system of claim 102, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of fixed-sized blocks of a main text of the plurality of e-mail messages.

105. (New) The system of claim 102, wherein the generating hash values includes:

performing a plurality of hashes on a main text of the plurality of e-mail messages using a plurality of different hash functions.

106. (New) The system of claim 102, wherein the generating hash values includes:

performing a plurality of hashes on a main text of the plurality of e-mail messages using a same hash function.

107. (New) The system of claim 102, wherein the generating hash values includes:

attempting to expand an attachment of the plurality of e-mail messages, and hashing the attachment after attempting to expand the attachment.

108. (New) The system of claim 102, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of variable-sized blocks of an attachment of the plurality of e-mail messages.

109. (New) The system of claim 102, wherein the generating hash values includes:

performing a plurality of hashes on a plurality of fixed-sized blocks of an attachment of the plurality of e-mail messages.

110. (New) The system of claim 102, wherein the generating hash values includes:

performing a plurality of hashes on an attachment of the plurality of e-mail messages using a plurality of different hash functions.

111. (New) The system of claim 102, wherein the generating hash values includes:

performing a plurality of hashes on an attachment of the plurality of e-mail messages using a same hash function.

112. (New) The system of claim 102, further comprising:

comparing the generated hash values to hash values corresponding to known unwanted e-mails.

113. (New) The system of claim 112, wherein the known unwanted e-mails include at least one of e-mails containing a virus, e-mails containing a worm, and unsolicited commercial e-mails.

114. (New) The system of claim 102, wherein the generating hash values includes:

hashing at least one of a main text and an attachment to generate one or more first hash values, and

hashing a concatenation of first and second header fields to generate a second hash value.

115. (New) The system of claim 114, wherein the first and second header fields include a From header field and a To header field.

116. (New) The system of claim 114, wherein the counting the number of the generated hash values corresponding to the message body associated with the one of the

plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages includes:

determining a first suspicion count based on a number of the hash values associated with the prior e-mail messages that match the one or more first hash values, and

determining a second suspicion count based on a number of the hash values associated with the prior e-mail messages that match the second hash value.

117. (New) The system of claim 116, wherein the determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message includes:

determining that the one of the plurality of e-mail messages is a potentially unwanted email message when the first suspicion count is higher than the second suspicion count.

118. (New) The system of claim 102, further comprising:

taking remedial action when the one of the plurality of e-mail messages is a potentially unwanted e-mail message, the taking remedial action including at least one of:

discarding the one of the plurality of e-mail messages,

bouncing the one of the plurality of e-mail messages,

marking the one of the plurality of e-mail messages with a warning,

subjecting the one of the plurality of e-mail messages to a virus or worm detection process,

creating a notification message, and

generating a suspicion score for the one of the plurality of e-mail messages and using the suspicion score to identify further processing for the one of the plurality of e-mail messages.

119. (New) The system of claim 102, further comprising:

generating a suspicion score for the plurality of e-mail messages based on a result of the counting of the number of the generated hash values corresponding to the message

body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages; and

taking remedial action when the one of the plurality of e-mail messages is a potentially unwanted e-mail message, the taking remedial action including:

determining whether a newly received e-mail message exceeds a mail quota,

identifying an earlier-received e-mail message with a highest suspicion score,

determining whether the suspicion score of the newly received e-mail message is lower than the suspicion score of the earlier-received e-mail message when the newly received e-mail message exceeds the mail quota,

deleting the earlier-received e-mail message when the suspicion score of the newly received e-mail message is lower than the suspicion score of the earlier-received e-mail message, and

storing the newly received e-mail message.

120. (New) The system of claim 102, wherein the generating hash values and the counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with prior e-mail messages are performed incrementally as the plurality of e-mail messages are being received.

121. (New) The system of claim 120, further comprising:

generating a suspicion score for the plurality of e-mail messages based on a result of the counting of the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages; and

taking remedial action when the suspicion score of an e-mail message of the plurality of e-mail messages is above a threshold, the taking remedial action including rejecting the e-mail message.

122. (New) The system of claim 121, wherein the rejecting occurs before the e-mail message is completely received.

123. (New) The system of claim 102, further comprising:

comparing the generated hash values to known legitimate mailing lists; and

passing the plurality of e-mail messages without further examination when the generated hash values match one or more of the known legitimate mailing lists.

124. (New) The system of claim 123, wherein the comparing the generated hash values includes:

determining whether the plurality of e-mail messages originated from the known legitimate mailing lists.

125. (New) The system of claim 102, wherein the generating hash values includes:

hashing a main text to generate a first hash value, and

hashing sender-related header fields to generate one or more second hash values.

126. (New) The system of claim 125, wherein the sender-related header fields include at least one of a From header field, a Sender header field, and a Reply-To header field.

127. (New) The system of claim 125, wherein the counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages includes:

determining a first suspicion count based on a number of the hash values associated with the prior e-mail messages that match the first hash value, and

determining one or more second suspicion counts based on a number of the hash values associated with the prior e-mail messages that match the one or more second hash values.

128. (New) The system of claim 127, wherein the determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message includes:

determining that the one of the plurality of e-mail messages is a potentially unwanted e-mail message when the first suspicion count is higher than the one or more second suspicion counts.

129. (New) The system of claim 102, wherein the generating hash values includes:

hashing a main text of the plurality of e-mail messages to generate a main text hash, and

hashing at least one header field of the plurality of e-mail messages to generate at least one header hash.

130. (New) The system of claim 129, wherein the counting the number of the generated hash values corresponding to the message body associated with the one of the plurality of e-mail messages that match the hash values corresponding to the message body associated with the prior e-mail messages includes:

determining whether the main text hash matches a higher number of the hash values associated with the prior e-mail messages than the at least one header hash; and

wherein the determining that one of the plurality of e-mail messages is a potentially unwanted e-mail message includes:

determining that the one of the plurality of e-mail messages is a potentially unwanted e-mail message when the main text hash matches a higher number of the hash values associated with the prior e-mail messages than the at least one header hash.

131. (New) The system of claim 102, wherein the processing the e-mail messages by removing the HTML comments and the HTML tags from the e-mail messages occurs before the generating the hash values.

132. (New) The system of claim 102, wherein the processing the e-mail messages by removing the HTML comments and the HTML tags from the e-mail messages occurs in parallel with the generating the hash values.

133. (New) The system of claim 102, wherein the portions of each message body of the plurality of e-mail messages that have been processed by removing the HTML comments and the HTML tags include blocks.